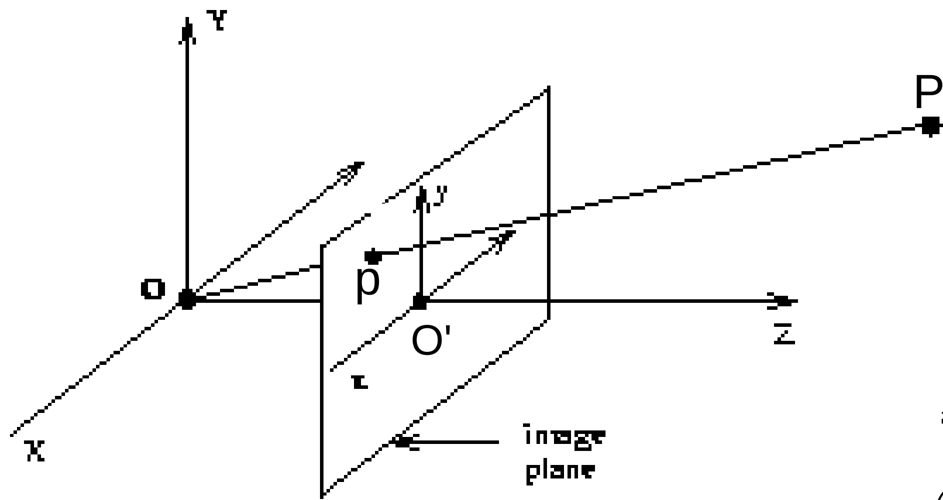
9º CAPÍTULO

Geometria Projectiva

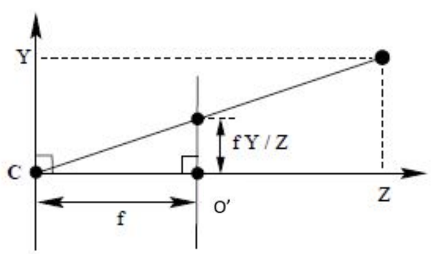
Prof. Arnaldo Abrantes / anotado por Prof. Nuno Pinho da Silva

- A projecção dum ponto 3D no plano de imagem pode ser descrita por uma projecção de perspectiva



$$P \leftarrow (X, Y, Z)^T$$
$$p \leftarrow (x, y)^T$$

$$x = f\,\frac{X}{Z} \qquad y = f\,\frac{Y}{Z}$$



Em coordenadas homogéneas:

$$\begin{pmatrix} \lambda x \\ \lambda y \\ \lambda \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

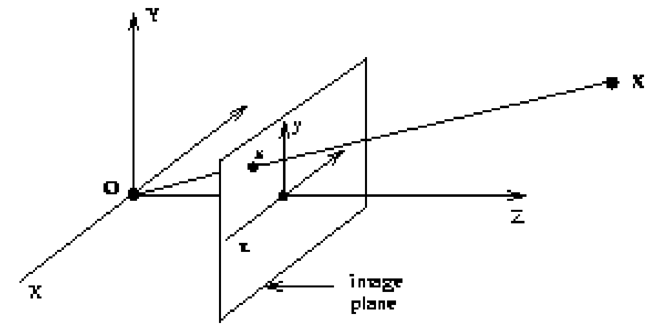
- ***Dificuldade***: Os pontos P no espaço 3D estão expressos em coordenadas métricas e usam o referencial da câmara (geralmente não se tem acesso a este sistema de coordenadas)

- ***Motivação***: As equações da projecção de perspectiva são não lineares quando estão expressas em coordenadas cartesianas, mas são lineares quando se exprimem em coordenadas homogéneas
  - Obs: Esta é uma característica de todas as transformações da geometria projectiva (e não apenas da projecção de perspectiva)

- ***Propriedades***: Coordenadas projectivas ou homogéneas
  - Dois vectores $(x_1, \cdots, x_n, x_{n+1})^T$ e $(\lambda x_1, \cdots, \lambda x_n, \lambda x_{n+1})^T$ representam o mesmo ponto para qualquer $\lambda \neq 0$
  - Um vector com coordenadas homogéneas $(x_1, \cdots, x_n, x_{n+1})^T$ representa o ponto com coordenadas cartesianas $\left(\frac{x_1}{x_{n+1}}, \cdots, \frac{x_n}{x_{n+1}}\right)^T$

- Sistemas de coordenadas envolvidos: câmara, imagem e mundo
  - 1) Referencial da câmara

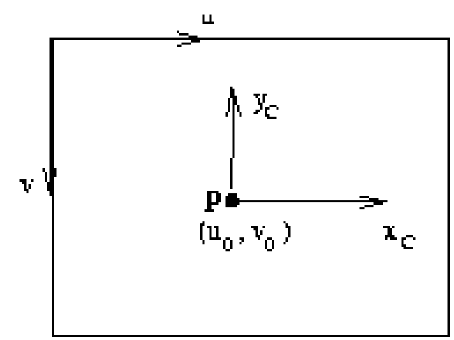
$$\lambda\begin{pmatrix} x^c \\ y^c \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} X^c \\ Y^c \\ Z^c \\ 1 \end{pmatrix}$$



  - 2) Referencial da imagem

$$\lambda\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} k_u & 0 & u_0 \\ 0 & -k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \lambda x^c \\ \lambda y^c \\ \lambda \end{pmatrix}$$

$$k_u x^c = u - u_0$$
$$k_v y^c = v_0 - v$$

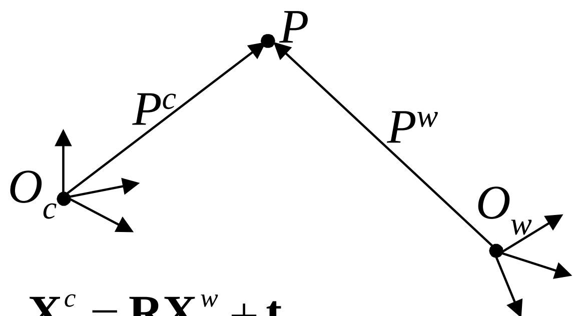


- Os parâmetros intrínsecos são 4 e caracterizam:
  - dimensão dos pixels (escalamentos vertical e horizontal, $fk_{u,}$ fk_v)
  - posição do ponto principal (u_0,v_0) - intersecção do eixo óptico com o plano de imagem

– 3) Referencial do mundo (transformação rígida)

- É mais prático expressar as coordenadas dos pontos 3D, usando um referencial do mundo {***w***} diferente do referencial da câmara {***c***}

$$P^c = R_w^c P^w + O_w^c$$



P^c - ponto P expresso em {***c***}

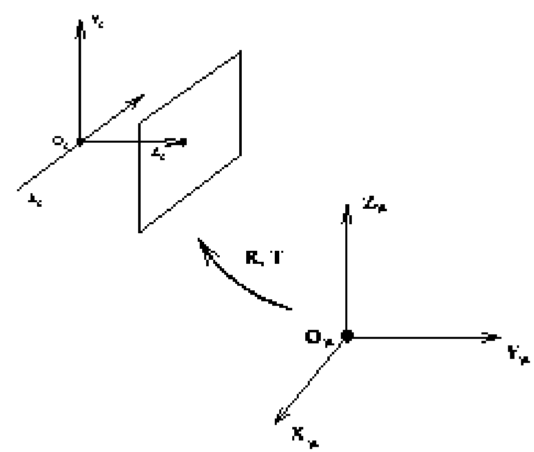
P^w - ponto P expresso em {***w***}

R_w^c - Rotação de {***w***} para {***c***}.
(referencial {***w***} expresso em {***c***})

O_w^c - origem de {***w***} expresso em {***c***}

$$\mathbf{X}^c = \mathbf{R}\mathbf{X}^w + \mathbf{t}$$

$$\lambda \begin{pmatrix} X^c \\ Y^c \\ Z^c \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} X^w \\ Y^w \\ Z^w \\ 1 \end{pmatrix}$$



- Os parâmetros extrínsecos são 6:
  – três parâmetros de rotação
  – três parâmetros de translação

- Concatenando os diferentes modelos, resulta:

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} k_u & 0 & u_0 \\ 0 & -k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} X^w \\ Y^w \\ Z^w \\ 1 \end{pmatrix}$$

- Ou mais simplesmente,

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} (\mathbf{R} \mid \mathbf{t}) \begin{pmatrix} X^w \\ Y^w \\ Z^w \\ 1 \end{pmatrix}$$

$$\alpha_u = f k_u$$

$$\alpha_v = -f k_v$$

- Finalmente….

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = P \begin{bmatrix} X^w \\ Y^w \\ Z^w \\ 1 \end{bmatrix}$$

$$\mathbf{P} = \mathbf{C}(\mathbf{R} \mid \mathbf{t})$$

$$\mathbf{C} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

Propriedades da matriz de projecção

- O espaço nulo de **P** é o centro óptico da câmara, expresso no referencial do mundo, ou seja $\mathbf{O}_c = -\mathbf{R}^T\mathbf{t}$

- Quando o referencial do mundo está alinhado com o referencial da câmara, então tem-se que $\mathbf{P} = \mathbf{C}(\mathbf{I} \mid \mathbf{0})$

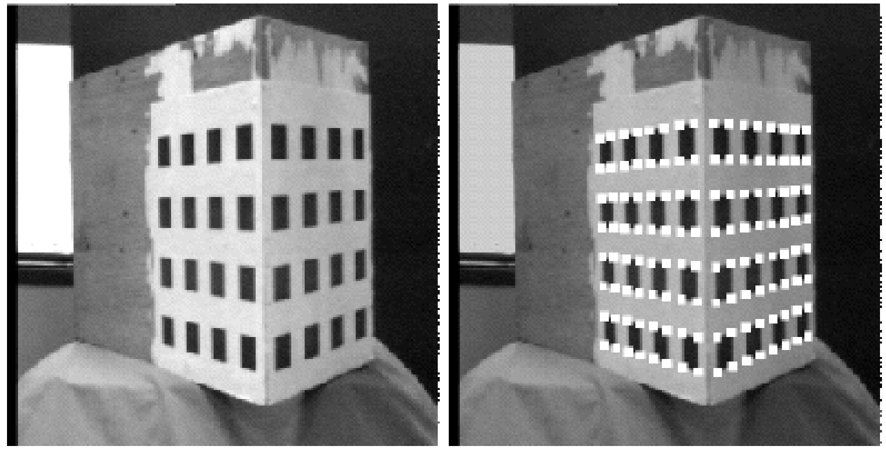
- Se as medidas na imagem forem *normalizadas*, então $\mathbf{P} = (\mathbf{I} \mid \mathbf{0})$, o que corresponde a usar pontos com coordenadas $\mathbf{x}_c = \mathbf{C}^{-1}\mathbf{x}_i$

- Um ponto **x** medido na imagem é retroprojectado no espaço 3D como sendo um raio com equação $\mathbf{X}(\lambda) = \mathbf{O}_c + \lambda(\mathbf{C}\mathbf{R})^{-1}\mathbf{x}$. Ou seja, todos os pontos **X** são projectados na imagem num mesmo ponto, **x**

- ***Objectivo da calibração***: determinar parâmetros intrínsecos (ou seja, a matriz **C**)
- ***Algoritmo***:
  - 1) Determinar a matriz **P**, usando um conjunto de *n* medidas na imagem, cujas coordenadas 3D no referencial do mundo são conhecidas

$$\mathbf{x} = \mathbf{P}\begin{pmatrix} \mathbf{X} \\ 1 \end{pmatrix}$$

$$\lambda \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}$$



  - 2) Decompôr a matriz **P** nas matrizes **C**, **R** e **t**. Por exemplo, através da decomposição RQ

$$\mathbf{P} = \mathbf{C}(\mathbf{R} \mid \mathbf{t})$$

– Cada correspondência dá origem a duas equações

$$x_i = \frac{p_{11}X_i + p_{12}Y_i + p_{13}Z_i + p_{14}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}} \qquad y_i = \frac{p_{21}X_i + p_{22}Y_i + p_{23}Z_i + p_{24}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}}$$

$$x_i\left(p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}\right) = p_{11}X_i + p_{12}Y_i + p_{13}Z_i + p_{14}$$

$$y_i\left(p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}\right) = p_{21}X_i + p_{22}Y_i + p_{23}Z_i + p_{24}$$

– Estas equações podem ser escritas em forma matricial,

$$\begin{pmatrix} X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -x_iX_i & -x_iY_i & -x_iZ_i & -x_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -y_iX_i & -y_iY_i & -y_iZ_i & -y_i \end{pmatrix}\mathbf{p} = \mathbf{0}$$

$$\mathbf{p} = \left(p_{11}, p_{12}, p_{13}, p_{14}, p_{21}, p_{22}, p_{23}, p_{24}, p_{31}, p_{32}, p_{33}, p_{34}\right)^T$$

– Usando $n \geq 6$ correspondências, obtém-se um sistema de *2n* equações homogéneas, com *11* incógnitas

At least 28 points should be used [Hartley & Zisserman]

$$\mathbf{Ap} = \mathbf{0}$$ (dimensão de **A** é *2n* x *11*)

Determinação da matriz **P** (cont.)

- Em geral, o sistema **Ap** = **0** não tem uma solução exacta
  - ***<u>Solução:</u>*** encontrar a solução que minimiza |**Ap**|, sujeita à restrição de |**p**|=1.
    - determinar o vector próprio de $\mathbf{A}^T\mathbf{A}$ com menor valor próprio, ou alternativamente
    - determinar o vector correspondente ao menor valor singular da decomposição SVD de **A**
- A solução *linear* encontrada pode ser melhorada de forma iterativa, através de utilização de método de optimização não linear

$$\min_{\mathbf{p}} \sum_i \left((x_i, y_i) - P(X_i, Y_i, Z_i, 1) \right)^2$$

- Método alternativo
  - Impôr $p_{34} = 1$ e obter solução de mínimos quadrados do sistema (agora não homogéneo) sobredeterminado $\mathbf{A}\mathbf{p}^{\#} = \mathbf{b}$

$\mathbf{p}^{\#} = \mathbf{A}^{+}\mathbf{b} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$ onde $\mathbf{p}^{\#} = \left(p_{11}, p_{12}, p_{13}, p_{14}, p_{21}, p_{22}, p_{23}, p_{24}, p_{31}, p_{32}, p_{33}\right)^T$

- Decomposição de **P** nas matrizes **C**, **R** e **t**
  - Seja **M** = **CR**.
    - 1) Factorizar **M** em **CR** usando, por exemplo, a decomposição matricial RQ (produto de matriz triangular superior por matriz de rotação). Obtém-se com este procedimento as matrizes **C** e **R**
    - 2) Calcular o vector translação, usando

$$\mathbf{t} = \mathbf{C}^{-1}\left(p_{14}, p_{24}, p_{34}\right)^T$$

  - Observação: Com este algoritmo, surge um parâmetro adicional, *k*, na matriz de parâmetros intrinsecos

$$\mathbf{C} = \begin{pmatrix} \alpha_u & k & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

Decomposição em valores singulares (SVD)

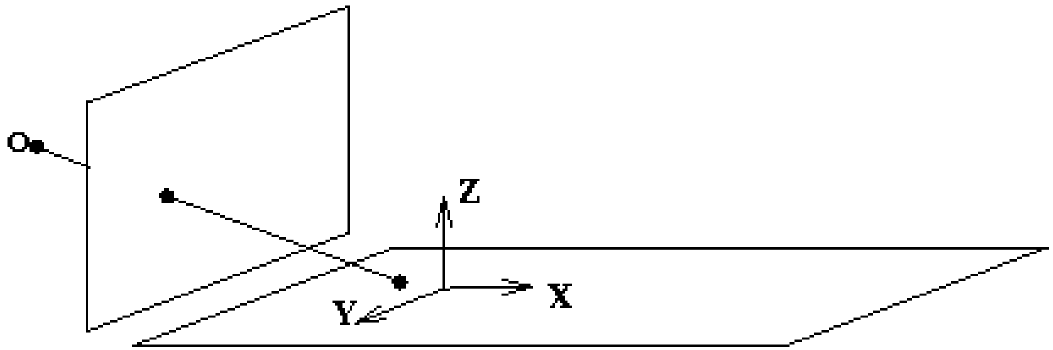
- ***Definição:*** Seja **A** uma matriz real de dimensão *m* x *n*. Então existem duas matrizes ortogonais **U** (*m* x *m*) e **V** (*n* x *n*), tais que

$$\mathbf{U}^T\mathbf{A}\mathbf{V} = diag(\sigma_1, \cdots, \sigma_p) \qquad p = \min\{m, n\}$$

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$$

- Os escalares σ_i designam-se por *valores singulares* da matriz **A**.
  - Estes valores são os comprimentos dos eixos do hiperelipsóide **A***x*, onde *x* é vector de norma unitária. A decomposição SVD pode portanto ser usada para encontrar as direcções que são mais “amplificadas” ou mais “atenuadas” pela multiplicação por **A**
  - A SVD é útil também para obter as bases ortonormadas que definem o espaço coberto por **A**, assim como o seu espaço nulo. Seja σ_r o menor valor singular não nulo da matriz **A**, então **A** tem característica (*rank*) *r*, o seu espaço coberto é definido pelas primeiras r colunas de **U**, $\{u_1, \cdots, u_r\}$ e o seu espaço nulo pelas últimas colunas de **V,** $\{v_{r+1}, \cdots, v_n\}$

- Se escolher o referencial do mundo de tal forma que os pontos no plano têm coordenadas **Z** nulas, então nesse caso resulta um modelo mais simples (8 parâmetros)



$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{14} \\ p_{21} & p_{22} & p_{24} \\ p_{31} & p_{32} & p_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$$

  – A matriz 3x3 representa uma transformação geral entre dois planos (plano 3D/imagem ou então imagem1/imagem2, resultante da observação de um conjunto de pontos pertencentes a um plano)

Homografia – Movimento é livre mas o mundo é plano

Mundo 3D
$\mathbf{X}_i$

$\mathbf{x}_i^1 = \mathbf{H}_1\mathbf{X}_i$

Imagem 1
$\mathbf{x}_i^1$

$\mathbf{x}_i^2 = \mathbf{H}_2\mathbf{X}_i$

Imagem 2
$\mathbf{x}_i^2$

$$\boxed{\mathbf{x}_i^2 = \mathbf{H}_{1,2}\mathbf{x}_i^1}$$

$$\mathbf{H}_{1,2} = \mathbf{H}_2\,\mathbf{H}_1^{-1}$$

(210, 56)

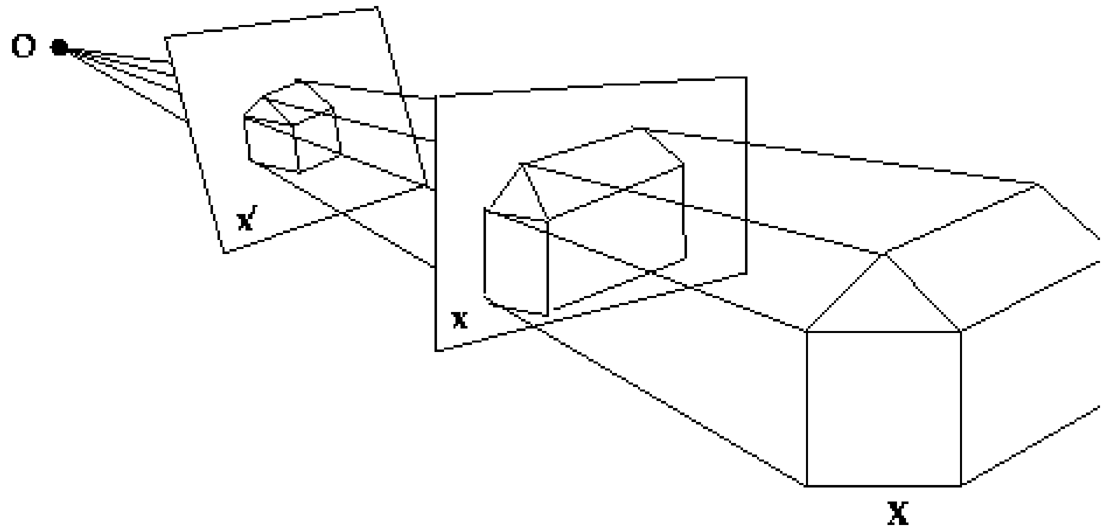
(114, 54)

- Quando a câmara se move, mas com o seu centro óptico mantido fixo, então os pontos em correspondência nas imagens estão relacionados através de transformações projectivas planas

$$\mathbf{x} = \mathbf{C}\left(\mathbf{I}|\mathbf{0}\right)\begin{pmatrix}\mathbf{X}\\1\end{pmatrix} = \mathbf{CX}$$

$$\mathbf{x}' = \mathbf{C}\left(\mathbf{R}|\mathbf{0}\right)\begin{pmatrix}\mathbf{X}\\1\end{pmatrix} = \mathbf{CRX}$$

$$= \mathbf{CRC}^{-1}\mathbf{x}$$



$$\mathbf{x}' = \mathbf{H}\mathbf{x} \qquad \mathbf{H} = \mathbf{CRC}^{-1}$$

  – ***Obs***: a matriz **H** não depende da estrutura 3D dos pontos observados

Exemplo: rotações sintéticas

Imagem original de corredor

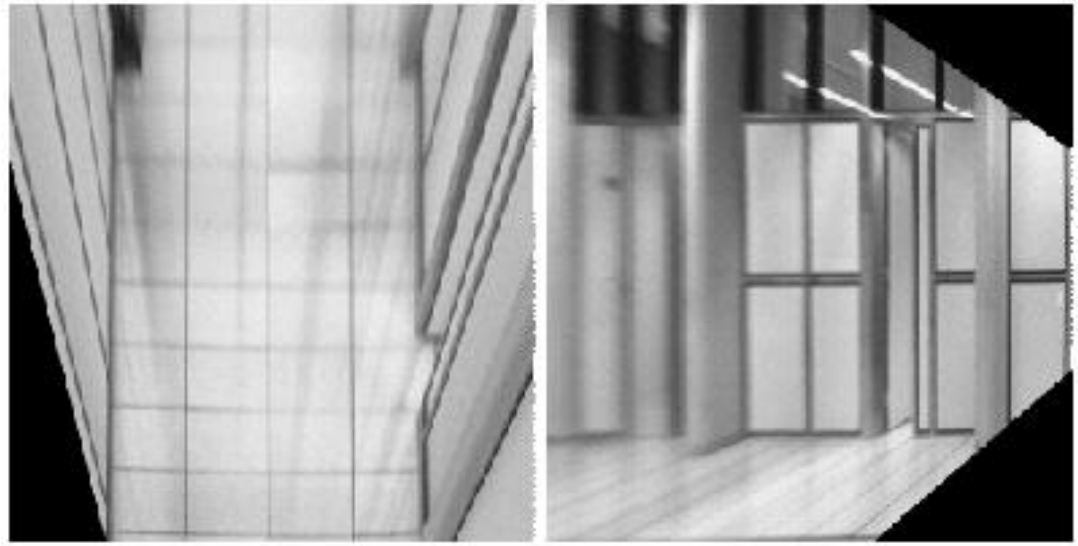
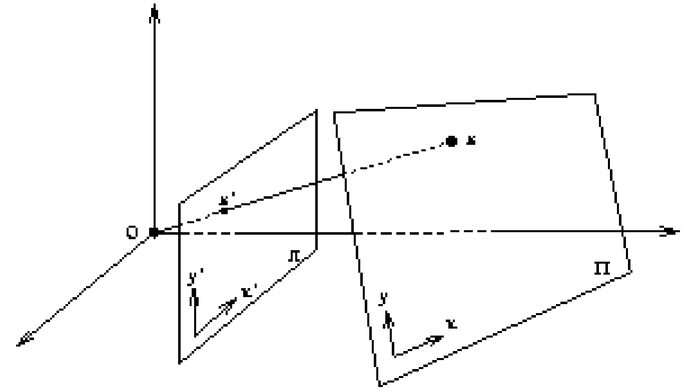


Imagem produzidas sinteticamente

- Ambas as imagens correspondem a rotações sintéticas da câmara, relativamente ao seu centro óptico
  - esquerda; transformação (*warping*) da imagem original por forma a se obterem ladrilhos quadrados
  - direita; transformação (*warping*) da imagem original por forma a se obter uma porta rectangular

– São necessários 4 correspondências para definir completamente uma transformação homográfica

$$\lambda\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$



$$\mathbf{x}' = \mathbf{Hx}$$

- Cada correspondência produz duas equações lineares nos elementos de **H**

$$x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}} \qquad y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}$$

$$x'\left(h_{31}x + h_{32}y + h_{33}\right) = h_{11}x + h_{12}y + h_{13}$$
$$y'\left(h_{31}x + h_{32}y + h_{33}\right) = h_{21}x + h_{22}y + h_{23}$$

– O oposto também se verifica. Através duma homografia é sempre possível transformar quaisquer 4 pontos em posições arbitrárias noutros quaisquer 4 pontos arbitrários.

- Possível solução:

  – Fazer h_{33}=1 e resolver sistema linear de 2*n* equações e 8 incógnitas

$$\mathbf{Ah} = \mathbf{b}$$

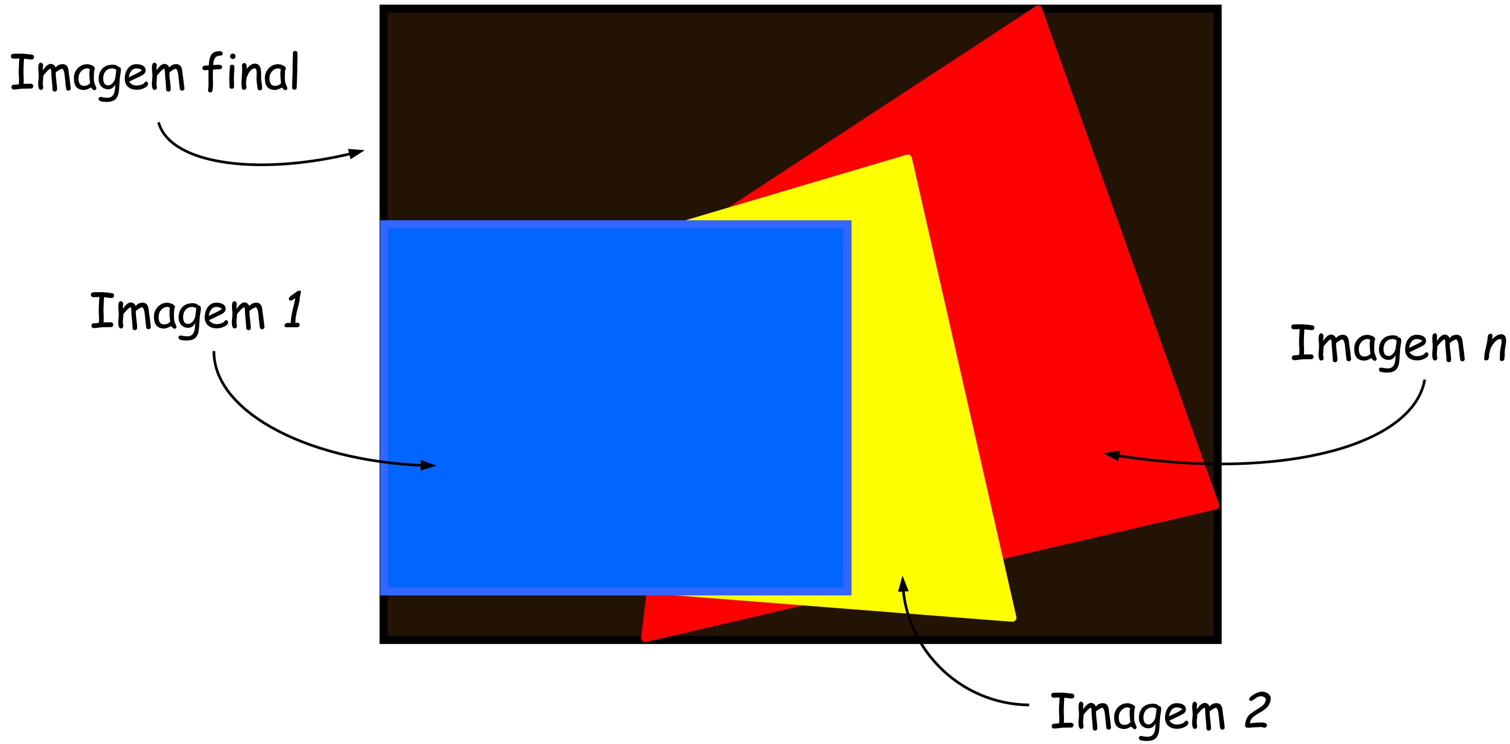
Exemplo com $n = 4$ ⟶

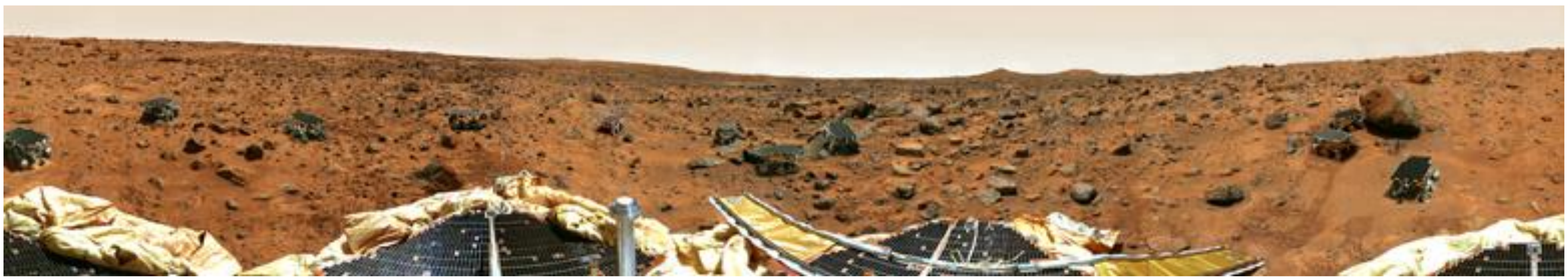
$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1'x_1 & -x_1'y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y_1'x_1 & -y_1'y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2'x_2 & -x_2'y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y_2'x_2 & -y_2'y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3'x_3 & -x_3'y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -y_3'x_3 & -y_3'y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4'x_4 & -x_4'y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -y_4'x_4 & -y_4'y_4 \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{pmatrix} = \begin{pmatrix} x_1' \\ y_1' \\ x_2' \\ y_2' \\ x_3' \\ y_3' \\ x_4' \\ y_4' \end{pmatrix}$$

  – Mínimos quadrados

$$\hat{\mathbf{h}} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b}$$

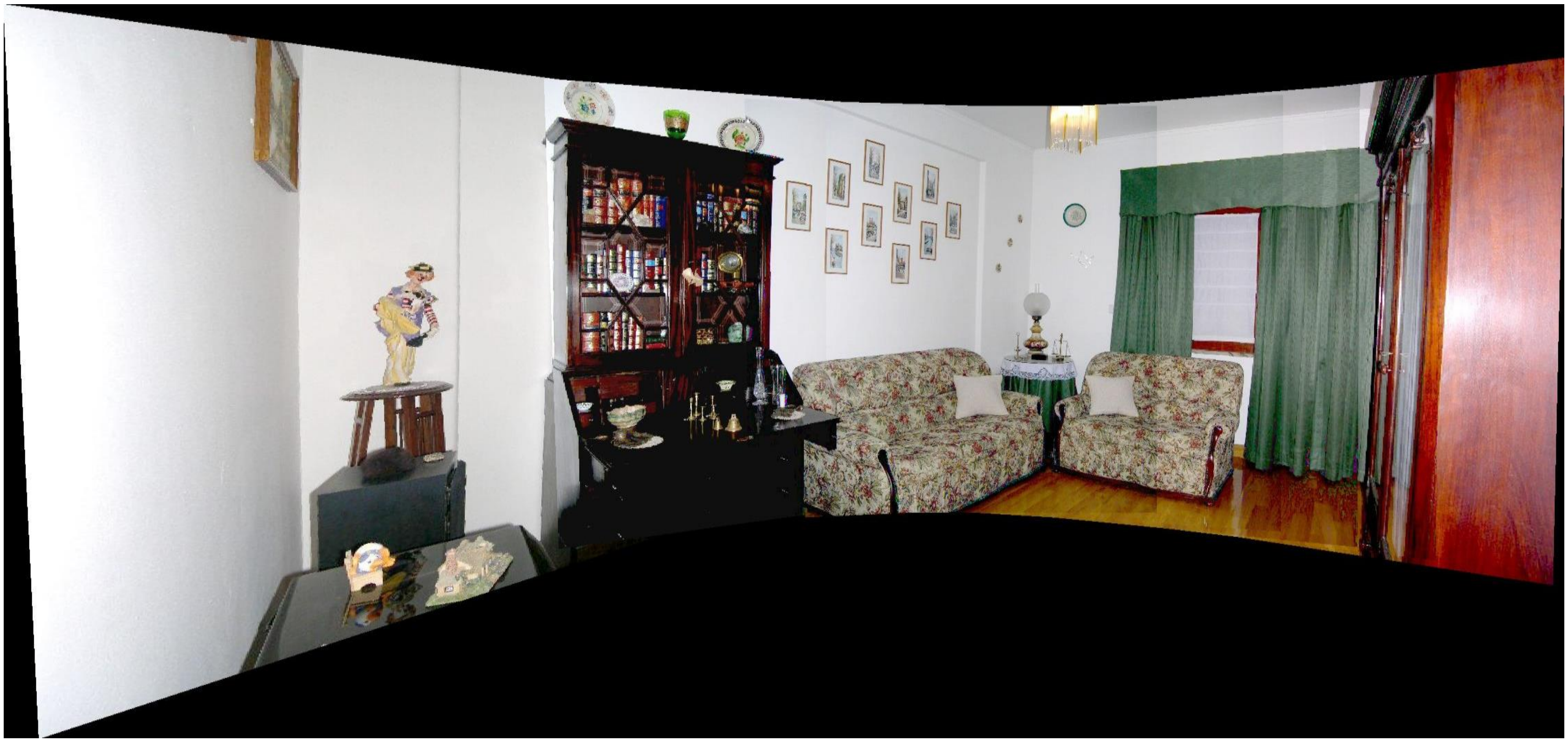
Aplicação na construção de mosaico





Mosaico da superfície de Marte

- Mosaico construído a partir de 8 imagens (matriz H estimada usando 10 correspondências, definidas manualmente)



Exemplo

- Mosaico construído a partir de 7 imagens (matriz H estimada usando 8 correspondências, definidas manualmente)

Edifício do DEETC